

1 DIVERSIFIED HOST BASED ROUTE SELECTION METRIC
2

3 Inventor(s): Michael T. Roeder

4 Mark Gooch
5
67 TECHNICAL FIELD

8 Embodiments of the present invention relate generally
9 to communication networks, and more particularly to a
10 diversified host based route selection metric.

11
12 BACKGROUND

13 It is desirable to implement routing systems that
14 search network route caches to find the forwarding
15 information for packets to be routed rather than searching
16 a complete forwarding information base, FIB, that contains
17 the forwarding information for all active routes for that
18 purpose; this invention provides a mechanism for
19 efficiently distributing the routing load between a network
20 route cache and a host route cache (or optionally a FIB.)
21 It is desirable to distribute the routing load between a
22 network route cache and a host route cache efficiently
23 because it reduces the number of entries that must be
24 supported in those caches.

1 A router is any device that is capable of routing
2 information between networks using a FIB, that it uses to
3 find the forwarding information for packets destined for
4 all known networks. A known network or host may be defined
5 as a network or host for which there is a route in the FIB
6 that has valid forwarding information.

7 A router may employ a variety of mechanisms to route
8 the packets. The router may route packets by searching the
9 FIB for a route to the destination of the packets and then
10 forwarding the packet with the forwarding information from
11 the matching route in the FIB. In addition to searching
12 the FIB, a router may employ caching mechanisms in hardware
13 or in software that hold forwarding information for
14 frequently accessed destinations. These caching mechanisms
15 may employ any combination of a host route cache and a
16 network route cache, at any one time the caches are capable
17 of routing to only a subset of the destinations that are
18 accessible using the information from the FIB. Both of
19 these caches are searched with the destination internet
20 protocol, IP, address of the packet to be routed. The
21 entries in the host route cache will match at most one
22 address, while the entries in a network route cache may
23 match one or more addresses. The entries in the network
24 route cache and in the host route cache will contain

1 forwarding information that is to be used for packets that
2 are routed using a matching cache entry. The advantage of
3 using these caches is that it is less expensive to build a
4 high performance routing cache than it is to build a high
5 performance FIB, as a result cost savings can be realized
6 by using a set of high performance caches (network and host
7 route caches) and a lower performance FIB in place of a
8 single high performance (and expensive) FIB.

9 The network route cache is used to provide subnet
10 based route lookup functionality. The network route cache
11 is searched using the destination address of a packet to be
12 routed and if the cache contains a route to the destination
13 address of the packet it will route the packet using the
14 forwarding information from the matching route. Thus, the
15 network route cache may be used as a cache to store a
16 subset of all of the routes that are known to the router.

17 It is desirable to use a combination of a network
18 route cache and host route cache efficiently such that the
19 size of these two caches may be limited for a reduction in
20 the cost of devices that employ these two caches.

21 Therefore, there is need in the current technology for new
22 methods and techniques that will effectively and
23 efficiently use the limited amount of space in a network
24 route cache.

1 Using present methods many routers store all of the
2 active entries from their RIB, routing information base, in
3 a high performance FIB that is used to make the majority of
4 the routing decisions, and this is very expensive. These
5 routers do not route packets using a network route cache,
6 these routers route using the full set of active routes in
7 the RIB along with optionally a host route cache. This
8 requires a very large and expensive forwarding table in the
9 high performance routing path. Embodiments of the
10 invention allow routers to efficiently manage a network
11 route cache; alternate mechanisms for maintaining such a
12 cache (for example LRU [Least Recently Used] or random
13 replacement) do not lead to efficient cache utilization.

14 Therefore, the current technology is limited in its
15 capabilities and suffers from at least the above
16 constraints.

17

1 SUMMARY OF EMBODIMENTS OF THE INVENTION

2 An embodiment of the invention ranks a set of routes
3 in a network route cache in order of value. For the
4 purposes of discussion in this disclosure the "value" of a
5 route will be defined to be a function of the "breadth" and
6 "frequency" of that route, where the function is an
7 embodiment of the invention. The term "frequency" or
8 "frequency of use" for a route refers to the number of
9 packets per unit time that were routed using that route.
10 The term "breadth" or "breadth of use" for a route refers
11 to the number of distinct destination hosts that that route
12 was used to route packets to per unit time. The value of a
13 route can be used to determine whether or not that route
14 should be stored in a network route cache, or if the
15 individual destinations within that route should be stored
16 in the host route cache on an as needed basis.

17 An embodiment of the invention allows the use of
18 network routing caches to be optimized such that routes
19 that are not used to route to a large number of
20 destinations can be offloaded into a less expensive host
21 route cache.

22 An embodiment of this invention includes two
23 components - a profiler and an analyzer. In an embodiment
24 of a profiler, a method for a diversified host based route

1 selection metric, includes: using an array indexed by a
2 value, hash(Pkt.daIP), and written to with a packet
3 destination address, Pkt.daIP, to profile the width and
4 breadth of use for routes in a routing table.

5 In an embodiment of an analyzer, a method of
6 evaluating the value, includes: configuring the profiler to
7 profile subsets of routes in the network route cache and
8 analyzing the profiler output for each subset of routes in
9 the network route cache to determine the relative value of
10 each route in the network route cache.

11 These and other features of an embodiment of the
12 present invention will be readily apparent to persons of
13 ordinary skill in the art upon reading the entirety of this
14 disclosure, which includes the accompanying drawings and
15 claims.

16

1 BRIEF DESCRIPTION OF THE DRAWINGS

2 Non-limiting and non-exhaustive embodiments of the
3 present invention are described with reference to the
4 following figures, wherein like reference numerals refer to
5 like parts throughout the various views unless otherwise
6 specified.

7

8 Figure 1 is a block diagram of a router that can
9 implement an embodiment of the invention.

10 Figure 2 is a block diagram of a route profiler, in
11 accordance with an embodiment of the invention.

12 Figure 3 is a flowchart of the method of a profiler in
13 accordance with an embodiment of the invention.

14 Figure 4 is a block diagram of an analyzer, in
15 accordance with an embodiment of the invention.

16 Figure 5 is a flowchart of a method of an analyzer in
17 accordance with an embodiment of the invention.

18

19

1 DETAILED DESCRIPTION OF PREFERRED EMBODIMENTS

2 In the description herein, numerous specific details
3 are provided, such as examples of components and/or
4 methods, to provide a thorough understanding of embodiments
5 of the invention. One skilled in the relevant art will
6 recognize, however, that an embodiment of the invention can
7 be practiced without one or more of the specific details,
8 or with other apparatus, systems, methods, components,
9 materials, parts, and/or the like. In other instances,
10 well-known structures, materials, or operations are not
11 shown or described in detail to avoid obscuring aspects of
12 embodiments of the invention.

13

14 WHY IT IS USEFUL TO KNOW THE BREADTH OF USE AND FREQUENCY
15 OF USE:

16 Figure 1 is a block diagram of a router 100 that can
17 implement an embodiment of the invention. The router 100
18 can receive and route packets 105 using the routing stack
19 115. During the course of normal routing activity, packets
20 105 are received and routed by the routing stack 115.
21 Typically the route selection algorithm 150 will program
22 network routes into the network route cache 106 and host
23 routes into the host route cache 123 as packets 105 are
24 received. The route selection algorithm does this in order

1 to reduce the number of lookups that are made into the FIB
2 122. The routing stack 115 is typically a combination of
3 hardware and software that routes packets 105 toward the
4 destination host specified as the destination address 110
5 in the packet. A router 100 uses a FIB and optionally some
6 combination of a host route cache 123 and a network route
7 cache 106 to find the appropriate forwarding information
8 120 to use to forward a packet 105. The forwarding
9 information 120 specifies the path through the network that
10 the packet 105 should follow and the operations that should
11 be performed to alter the packet 105 before it is forwarded
12 out of the system (these operations for IP routing over
13 Ethernet include changing the source and destination MAC
14 addresses of the packet.)

15 The network route cache 106 is a data structure that
16 is accessed by the routing stack 115 to find the forwarding
17 information 120 that should be used to forward packets 105.
18 When a packet 105 is received, the network route cache 106
19 is searched with the destination address 110 of the
20 incoming packet 105 to be routed. If the network route
21 cache 106 contains a route to a particular destination, it
22 will return to the routing stack 115 the forwarding
23 information 120 for that destination address 110; otherwise

1 the forwarding information will have to be found in the FIB
2 122 or host route cache 123 if one is present.

3 The host route cache 123 serves the same function as
4 the network route cache 106; it is searched with the
5 destination address 110 of the packet 105 and returns the
6 forwarding information 120 for that packet 105. The
7 difference between the network route cache 106 and the host
8 route cache 123 is that an entry in the host route cache
9 123 matches one and only one destination address while a
10 network route can be configured so that all destination
11 hosts within a particular subnet will match that route.

12 The network route cache 106 can be implemented by any
13 suitable structure such as a content addressable memory,
14 linked list, tree structure, or any structure that can be
15 searched for network routes that match destination
16 addresses. The host route cache 123 can be implemented by
17 any suitable structure such as a hash table, a content
18 addressable memory, a tree, or any structure that can be
19 searched for host routes that match destination addresses.
20 Various additional known components and modules that permit
21 the routing of a packet 105 are not shown in Figure 1 to
22 avoid obscuring the intent of embodiments of the invention.

23 An embodiment of the invention includes a profiler 135
24 and an analyzer 140; these elements provide the router 100

1 with data that a route selection algorithm 150 can use to
2 optimize the use of the route caches 123 and 106. Unlike
3 previous route profilers, a route profiler 135 in
4 accordance with an embodiment of the invention provides a
5 mechanism for measuring the relative value (where value is
6 defined to be a function of frequency and breadth as
7 described above) of all routes in the network route cache
8 106.

9 The ability to measure the relative value of a network
10 route is useful in a router 100 that uses a combination of
11 a network route cache 106 and a host route cache 123. The
12 data collected by the profiler 135 is used by the analyzer
13 140 to rank the routes in the network route cache 106 in
14 order of value. The ranking information that is generated
15 by the analyzer 140 is then used by a route selection
16 algorithm 150 to determine which routes in the routing
17 table should be programmed into the network route cache
18 106; destinations that need to be routed to using a route
19 that the route selection algorithm 150 does not place into
20 the network route cache 106 will be routed using either
21 host routes in the host route cache 123 or using the FIB
22 122.

23 The combination of the profiler 135 and the analyzer
24 140 allows the route selection algorithm 150 to select

1 routes for a network route cache 106 so that the
2 combination of a network route cache 106 and a host route
3 cache 123 can be efficiently utilized such that the size of
4 these two caches 106 and 123 may be reduced. Further the
5 profiler 135 provides the analyzer 140 with sufficient
6 information such that routing latency can be reduced when a
7 route is removed from the network route cache 106 by the
8 route selection algorithm 150. As a result of these two
9 properties a router 100 with a profiler 135 and analyzer
10 140 may achieve high performance routing with significant
11 cost savings by using a relatively small network route
12 cache 106 and a relatively small host route cache 123.

13 Periodically, the analyzer 140 will analyze all
14 entries in the profile table, gather statistics about those
15 entries, and optionally record the set of IP addresses
16 found in the profile table. This analysis may involve
17 classifying the set of routes in the network route cache
18 into subsets based upon value, as described in detail
19 below. In this embodiment of an analyzer 140, the set of
20 routes H will be split into N subsets H_i that will be
21 profiled individually. Over time the analyzer 140 will
22 move routes between these subsets such that the most
23 valuable routes are stored in lower numbered subsets. Then
24 when space is needed in the network route cache, routes in

1 the lower valued subsets will be replaced before the routes
2 in the higher valued subsets. The analysis performed by
3 the analyzer 140 will generate a ranking for every route in
4 the network route cache 106. (The routes will be ranked in
5 order of their value.)

6 Based upon the ranking information generated by the
7 analyzer 140 the route selection algorithm 150 will
8 determine if a route should be placed in the network route
9 cache 106 or if host routes for the destinations that this
10 route has been used to route to recently should be
11 programmed into the host route cache 123. In this
12 embodiment the route profiler 135 keeps a record of the
13 destination addresses that are in use within each route.
14 This usage information may be used to program the host
15 route cache 123 with entries to replace a network route in
16 the network route cache 106. This way the forwarding
17 decision latency is not increased significantly during the
18 transition of the route from the network route cache 106 to
19 the host route cache 123.

20 The route profiler 135 provides adequate information
21 to select a subset of routes, H, from a set of all
22 accessible routes in the FIB 122, S, such that when the
23 subset of routes, H, is placed in the network route cache
24 106 (Figure 1), the routing latency is reduced and the

1 efficiency of any combination of a network route cache 106
2 and a host route cache 123 is improved. The route profiler
3 135 enables this improvement in efficiency by providing a
4 measurement of the breadth and frequency of use for routes
5 that are profiled; this measurement can be used to select
6 the routes that should be programmed into the network route
7 cache 106. As noted above, an accessible host is defined
8 as a host that can be reached by a data packet 105 through
9 the router 100 using a route known to the router 100. This
10 known route is defined as an accessible route S.

11 The profiler 135 will typically be implemented using a
12 hash table; however the method may use a list, instead of a
13 hash table, for profiling route breadth and frequency. The
14 method of the profiler 135 typically uses the destination
15 address 110 of routed packets 105 as the sole input;
16 however it may also use different inputs to the profiler
17 135 other than the destination address 110 of packets 105.
18 These different inputs could possibly be used to measure
19 the breadth and frequency of flows [the combination of
20 source and destination information from a packet] as well
21 as host routes. An embodiment of the profiler 135 may be
22 designed so that the profiler 135 can be configured to
23 profile only selected routes in the network route cache
24 106. The method of identifying routes to be profiled may

1 be any suitable route identification method. The analyzer
2 140 that is used may also be varied.

3 An embodiment of the route profiler 135 can operate
4 either as a hardware device or as a software algorithm and
5 is used with a router 100 that routes with any suitable
6 routing protocol that permits routing between networks by
7 use of a routing table.

8 The route selection algorithm 150 will typically use
9 this ranking information to select routes that should be
10 removed from the network route cache 106 when existing
11 routes need to be removed to make room for new routes. The
12 route selection algorithm 150 will typically remove from
13 the network route cache 106 the routes with the lowest rank
14 first.

15 Typically the route selection algorithm 150 will
16 remove one of the least valued routes when it needs to make
17 room for a new route. When a new route is added the
18 selection algorithm 150 will typically add that route to a
19 subset H_i that is between the most valued, H_0 , and least
20 valued subset H_N ; this subset will often be H_P where P is
21 $N/2$.

22 During the course of normal routing activity packets
23 105 are received and routed by the routing stack 115.

24 Typically the route selection algorithm 150 will program

1 network routes into the network route cache 106 and host
2 routes into the host route cache 123 as packets 105 are
3 received. The route selection algorithm 150 does this in
4 order to reduce the number of lookups that are made into
5 the FIB 122.

6

7 Figure 2 is a block diagram of a route profiler 135,
8 in accordance with an embodiment of the invention. An
9 embodiment of the route profiler 135 generates information
10 that the analyzer 140 can use to rank routes in the network
11 route cache 106 in order of value. In an embodiment, the
12 route profiler 135 typically includes the following
13 elements as discussed below and operates with the network
14 route cache 106, host route cache 123, and analyzer 140.

15 What follows is a brief description of the profiler
16 components: 210, 230 and 220. The operation of these
17 components 210, 230, and 220 will be described in greater
18 detail with Figure 4.

19 The profile control table 210 is used by the analyzer
20 140 to select a subset of the routes in the network route
21 cache 106 to profile. Typically the router 100 would not
22 profile all of the routes in the network route cache 106
23 simultaneously; the subset that is to be profiled is
24 selected by programming the profile control table 210. In

1 an embodiment, the profile control table 210 has one (1)
2 profile bit 215 per entry in the network route cache 106.
3 If a profile bit 215 is set in the profile control table
4 210 for a corresponding entry in the network route cache
5 106, then this means that profiling is enabled for the
6 corresponding network route cache entry. The profile
7 control table 210 will typically be programmed by the
8 analyzer 140.

9 The hash operation performed by the module 230 is used
10 to index the profile table 220. This hash operation 230
11 typically involves the use of a hash key 226 that can be
12 programmed to vary the hash function that is used. The
13 programmable hash operation performed by the module 230 is
14 used to take the destination address from a packet that is
15 being routed and generate an index into the profile table
16 220. As known to those skilled in the art, hashing is the
17 transformation a set of bits, or any numerically represent
18 able value, into a usually smaller fixed-length value or
19 address that represents the original value. It is noted
20 that it is within the scope of embodiments of the invention
21 to use all suitable hash functions. For example, this set
22 of suitable hash functions includes the function, $\text{hash}(\text{DA})$
23 $= \text{DA}$.

1 The profile table 220 is typically a random access memory that is addressed with the
2 hash value specified by the hash operation 230 and is written to with information about a routed
3 packet. The write operation is controlled with a write enable bit 215 that is typically supplied from
4 the profile control table 210. For this embodiment the information that is written to the hash table
5 for each packet that is routed using a profiled route in the network route cache 106 is the
6 destination address of the packet and the match location M of the route in the network route
7 cache that was used to route the packet. The match location M uniquely identifies the route in
8 the network route cache that was used to route the packet.

9 This write operation to the profile table 220 will
10 overwrite any existing data at the calculated particular
11 hash address 225, and as a result of this write operation,
12 the profiling operation should be viewed as a sampling
13 process. It is, therefore, typically necessary to perform
14 a profiling operation repeatedly using a different hash key
15 226 for each cycle of the profiling operation, this is done
16 to gain a more accurate picture of the value of a route in
17 the network route cache.

18 Periodically, an analyzer 140, will analyze all
19 entries in the profile table 220, gather statistics about
20 those entries, and optionally record the set of IP
21 addresses found in the profile table. The purpose of the
22 analyzer 140 is to rank the routes in the network route
23 cache in order of value such that the route selection
24 algorithm can intelligently decide which routes to leave in

1 the cache and which to remove. The analyzer 140 will be
2 discussed in greater detail below.

3 It is noted that the network route cache 106, profile
4 control table 210, and profile table 220 include functions
5 with logic circuitry and/or software that permits read and
6 write operations to the network route cache 106, Profile
7 Control Table 210, and Profile Table 220.

8

9 Figure 3 is a flowchart of a method 400 in accordance
10 with an embodiment of the invention. A determination is
11 made (405) if there is a route in a network route cache
12 106, M, that should be used to route a received packet 105.
13 If a route (in the network route cache) is used to route
14 the received packet 105, then a determination is made (410)
15 if the route M is a profiled route (i.e., a route that is
16 marked for profiling in the profile control table 210).

17 If the packet 105 is routed using a profiled route
18 from the network route cache 106 then information will be
19 written in the profile table in response to the packet 105
20 at a location specified by the hash operation 230. In an
21 embodiment of the invention the information that will be
22 written into the profile table 220 will be the destination
23 address 110 of the packet and the identifier, M, of the

1 network route in the network route cache 106 that was used
2 to route the packet.

3 This write operation allows the profiler 135 to build
4 up a pseudo random sample of destination addresses from
5 packets that were routed using profiled routes. An
6 analyzer 140 can use this sample of destination addresses
7 to determine the value of each profiled route relative to
8 the value of all the other profiled routes. The analyzer
9 140 can use this information to rank routes in the network
10 route cache in order of value.

11 As an example, an IP destination address 110 of an
12 incoming packet 105 is at 1.2.3.4, with a match location M
13 where M=23, and the hash operation address 225 is at
14 address=123. If the profile bit 215 has a value of 1 to
15 indicate profiling of the route, then the destination IP
16 address 110 and the match location M are written into the
17 profile table 220 at address=123. The destination address
18 110 of the next received packet 105 is then analyzed in the
19 same manner.

20 It is noted that multiple profile control tables 210
21 and multiple associated profile tables 220 can be used in
22 Figure 2, in order to simultaneously perform the profiling
23 of multiple subsets of routes.

24

1 Figure 4 is a block diagram of an embodiment of an
2 analyzer 140. In an embodiment, the analyzer 140 typically
3 includes the following elements as discussed below and
4 operates with the profiler 135 and the route selection
5 algorithm 150. The elements of an embodiment of an
6 analyzer 140 include a set IP Set 530, a value counter
7 array 520, and an analyzer algorithm 510.

8 IP Set 530 stores a set of structures 531 that contain
9 an IP address 531a, a route identifier 531b and a count
10 value 531c. Each structure 531 in IP Set 530 is uniquely
11 identifiable by the IP address 531a field. No two distinct
12 structures 531 in IP Set 530 will have the same IP address
13 531a. Elements can be added or deleted from IP Set 530.
14 IP Set 530 can be searched by IP address, if an IP address
15 is present in the set then IP Set 530 can be queried for
16 the count value or the route identifier associated with
17 that IP address and the count or route identifier will be
18 returned. In addition if an IP address is present in IP
19 Set 530 then the count value or route identifier associated
20 with that IP address can be modified. Also IP Set 530 can
21 be read, for example, linearly such that each entry 531 is
22 read one after another. However, in practice the IP Set
23 530 can be read by any suitable order, as the IP Set 530 is
24 typically data stored in memory (e.g., Random Access

1 Memory). IP Set 530 is used to keep a record of all of the
2 IP addresses that have been seen in the profiler table 220
3 along with their associated frequency count and the route
4 identifier of the route that was used to route to that
5 address.

6 The value counter array 520 represents an array of
7 counter structures 521. Each counter structure in the
8 array 521 corresponds to a particular route in the network
9 route cache; entry x in the value counter array 520
10 corresponds to entry x in the network route cache. Each
11 counter structure 521 is formed by a subset identifier 521b
12 and a breadth count 521c, together these reflect the
13 counters "value" as measured by the profiler 135 and
14 analyzer 140. The value of value counters can be compared;
15 two comparison methods are below:

- 16 1.) A value counter V1 is less valuable than
17 another value counter V2 if (V1.subset >
18 V2.subset) or if ((V1.subset == V2.subset)
19 and (V1.breadth < V2.breadth)).
- 20 2.) Two value counters, V1 and V2, are equivalent
21 if ((V1.subset == V2.subset) and (V1.breadth
22 == V2.breadth)).

23 The analyzer algorithm 510 periodically collects data
24 from the profiler 135 and recalculates the IP Set 530 and

1 value counters 520 based upon the contents of the profile
2 table 220 in the profiler 135. The method that one
3 embodiment of an analyzer algorithm 510 uses to generate
4 the IP Set 530 and value counters 520 is described in
5 figure 5.

6

7 Figure 5 is a flowchart of a method 600 in accordance
8 with an embodiment of the invention.

9 The analyzer method 600 will typically begin executing
10 after the router 100 begins initializing during boot and
11 before the route selection algorithm 150 has programmed any
12 routes into the network route cache 106. At this time the
13 network route cache will be empty and as a result all of
14 the N subsets H_i will also be empty. As execution of the
15 method 600 progresses the route selection algorithm 150
16 will add and remove routes from the network route cache.
17 When the route selection algorithm 150 adds a new route to
18 the network route cache it must assign that route to a
19 subset H_i . During execution of the method 600 routes will
20 be moved between subsets.

21 The analyzer will begin execution at step 610 and
22 continue on to step 620. At step 620 the method sets value
23 j of method 600 to 0. At step 630 the analyzer programs
24 the profiler to profile a subset, H_j , of routes in the

1 routing table and then pauses a period of time T1 (the
2 sampling period) while the profiler profiles a subset of
3 routes Hj.

4 Following step 630 the analyzer will gather the
5 information collected in the profiler during step 630. For
6 each entry in the profiler table 220 the analyzer will
7 check IP Set 530 to determine if the address, "A1", for
8 that entry appears in IP Set 530; if there is no entry in
9 IP Set with the address "A1" it will create one using
10 address A1 for the address 531a, the route identifier from
11 the profile table for the route identifier 531b and
12 MAX_COUNT for the count value 531c; if there already is an
13 entry in IP Set for the address "A1" then it will overwrite
14 the existing entry's count value with MAX_COUNT and will
15 overwrite the route identifier 531b with the route
16 identifier from the entry in the profiler table 220. Once
17 the analyzer has taken all the IP addresses that appeared
18 in the profile table 220 and used them to update IP Set it
19 will turn off profiling for all routes and it will clear
20 the profiler table and proceed to step 650. A typical
21 value for MAX_COUNT will be two (2); a typical value for N
22 will be eight (8). MAX_COUNT is a value that is used to
23 specify the number of sampling periods that an IP address

1 that isn't found in the profile table 220 will be left in
2 IP Set 530.

3 For step 650 it will increment j by one, then in step
4 660 it will check to see if j specifies a valid subset Hj.
5 If j specifies a valid set it will jump to step 630,
6 otherwise it will continue on to step 670.

7 For step 670 the analyzer method 600 will take the
8 information gathered in IP Set 530 and use it to rank the
9 routes in the network route cache. Step 670 will begin by
10 generating a value counter for every route in the network
11 route cache based upon the information in IP Set 530. Step
12 670 will iterate through all of the entries in IP Set 530
13 and for each entry, E1, 531 it will add the count value
14 531c to the value count 521 of the value counter 521 that
15 is associated with the route identifier of the current
16 entry, E1. (ValueCountersArray[route identifier
17 531b].value += count 531c) After the value counters are
18 updated as described above the routes will be ranked in
19 order of value based upon their corresponding entries in
20 the value counters array 520. A route, "R1", in the
21 network route cache 106 will be considered more valuable
22 than another route, "R2", if their respective counter
23 values "C1" and "C2" are such that $C1 > C2$. This ranking
24 information will be provided to the route selection

1 algorithm 150 whenever it requests it. After the rankings
2 have been generated the analyzer method will decrement the
3 count value 531c for every entry in IP Set, every entry 531
4 that has a count value 531c of 0 after its count is
5 decremented will be deleted from IP Set 530.

6 Step 680 follows step 670, during step 680 the routes
7 in the network route cache will be repartitioned into
8 subsets H_j such that the all the subsets H_j are similar in
9 size and such that some of the most valuable routes (as
10 determined in step 670) from each subset are promoted into
11 a more valuable subset and such that some of the least
12 valuable routes (as determined in step 670) from each
13 subset are demoted into a less valuable subset. Following
14 step 680 the method 600 will loop back to step 620.

15 The method uses the set of addresses 110 that appear
16 in the profile table 220 over a sampling period T_1 to
17 measure the relative value of maintaining each route in a
18 subset of routes, H_i , in the network route cache 106. This
19 sampling period occurs during step 630.

20

21 EXAMPLE USAGE:

22 What follows is a brief example of the route profiler
23 135 and analyzer 140 in use. Suppose that a router has a
24 network route cache with 6 routes in it, R1 through R6.

1 Further suppose that these routes are initially divided
 2 into groups as shown in Table 1. Further suppose that the
 3 router has an embodiment of the route profiler as described
 4 in Figure 2; this embodiment has a profile table 220 with 8
 5 entries. Table 1 represents the state of the value
 6 counters array 520 before the analyzer 140 begins with step
 7 610.

8

Route #	Subset	Value
1	0	0
2	0	0
3	0	0
4	1	0
5	1	0
6	1	0

9 Table 1, Value Counters

10

11 The router uses an embodiment of an analyzer 140 as
 12 described by Figures 4 and 5. For this embodiment of an
 13 analyzer the MAX_COUNT will be 2 and N will be 2. The
 14 routes begin evenly distributed between 2 subsets as
 15 required by step 610 of the analyzer method. The analyzer
 16 then runs to step 630, during step 630 the profiler
 17 profiles subset 0, H0. Table 2 describes the contents of
 18 the profiler table 220 after step 630.

1

Entry #	Address	Route #
0		
1	A1	1
2	A7	1
3		
4	A2	3
5		
6	A6	6
7		

2 Table 2, Profiler Table

3

4 This Table 2 was generated during the sampling period
5 of step 630 while profiling subset 0. During this period
6 at least four packets were received - one destined for
7 address A1, one for A7, one for A2 and one for A6. The
8 hash function 230 translated these addresses, A1, A7, A2
9 and A6 into the hash values 1, 2, 4 and 6 respectively.
10 The analyzer then proceeds to analyze this sample in step
11 640. Table 3 represents the contents of IP Set 530 after
12 step 640.

13

Address	Route #	Count
A1	1	2
A7	1	2

1 Table 3, IP Set

2

3 After step 640 the analyzer runs through steps 650,
 4 660, 630 and 640 again, this time steps 630 and 640 apply
 5 to subset 1, H1. This time only one address is found in
 6 the profiler table, it is A2 and was routed to using route
 7 3. Table 4 represents the contents of IP Set 530 after
 8 step 640 has run a second time.

9

Address	Route #	Count
A1	1	2
A7	1	2
A2	3	2
A6	6	2

10 Table 4, IP Set

11 After step 640 the analyzer runs through step 650, 660
 12 and 670. Step 670 cleared the value field for every route
 13 in the value counters array and then the count value from
 14 each address found in the IP set 530 was taken and added to
 15 the value for that addresses associated route in the value
 16 counters array 520 during step 670. Table 5 represents the
 17 value counter array 520 after executing step 670 for the
 18 first time.

Route #	Subset	Value
---------	--------	-------

1	0	4
2	0	0
3	0	2
4	1	0
5	1	0
6	1	2

1 Table 5, Value Counters

2

3 After step 670 the route subset assignments will be
4 adjusted based upon each routes rank within their current
5 subset. In this example the analyzer will swap two routes
6 between subsets 0 and 1. Since route 2 is the least
7 valuable member of subset 0 the analyzer moves route 2 from
8 subset 0 to a less valuable subset, subset 1. Since route
9 6 is the most valuable member of subset 1 the analyzer
10 moves route 6 from subset 1 to a more valuable subset,
11 subset 0. The value counters array 520 subset field 521b
12 will not be adjusted to reflect the changed subset
13 assignments until the next time step 670 runs. However
14 step 630 will use the new subset assignments generated
15 during step 680 immediately.

16 Once step 680 completes steps 620 through 660 are
17 rerun for both subsets. Table 6 describes the contents of
18 IP Set After steps 630 and 640 have been rerun for both
19 subsets 0 and 1.

1

Address	Route #	Count
A1	1	2
A7	1	1
A2	3	2
A6	6	1
A8	1	2
A4	2	2

2 Table 6, IP Set

3

4 During step 640 the analyzer 140 found one instance of
5 each of the following addresses: A1, A2, A8 and A4. No
6 instances of the addresses A7 or A6 were found in the
7 profiler table 220 and as a result the count values for
8 those addresses in IP Set 530 were decremented.

9 Table 7 reflects the value counters array 520 after
10 step 670 and 680 have each run again.

11

Route #	Subset	Value
1	0	5
2	1	2
3	0	2
4	1	0
5	1	0
6	0	1

12 Table 7, Value Counters

1

2 Based upon the value counters in Table 7 the analyzer
 3 chooses to swap two routes, route 2 and route 6. Route 2
 4 is moved to the higher valued subset, subset 0 because it
 5 is currently the most valuable member of subset 1. Route 6
 6 is moved to the lower valued subset, subset 1 because it is
 7 currently the least valuable member of subset 0.

8 Once step 680 completes a second time steps 620
 9 through 660 are rerun for both subsets. Table 8 describes
 10 the contents of IP Set After steps 630 and 640 have been
 11 rerun for both subsets 0 and 1.

12

Address	Route #	Count
A1	1	1
A2	3	2
A6	6	2
A8	1	1
A4	2	1
A5	5	2
A9	1	2
A3	4	2

13 Table 8, IP Set

14

15 During step 640 the analyzer 140 found one instance of
 16 each of the following addresses: A2, A6, A5, A9 and A3. No

1 instances of the addresses: A1, A7, A8 or A4 were found in
 2 the profiler table 220 and as a result the count values for
 3 those addresses in IP Set 530 were decremented. The count
 4 value for A7 dropped to 0 when it was decremented and as a
 5 result address A7 was removed from IP Set 530.

6 Table 9 reflects the value counters array 520 after
 7 step 670 and 680 have each run again.

8

Route #	Subset	Value
1	0	4
2	0	1
3	0	2
4	1	2
5	1	2
6	1	2

9 Table 9, Value Counters

10

11 Based upon the value counters in Table 9 the analyzer
 12 chooses to swap two routes, route 2 and route 4. Route 4
 13 is moved to the higher valued subset, subset 0 because it
 14 is currently one of the most valuable members of subset 1.
 15 Route 2 is moved to the lower valued subset, subset 1
 16 because it is currently the least valuable member of subset
 17 0.

1 The analyzer will continue to run in this fashion
2 until the router stops the analyzer task. This is an
3 example of the analyzer operation.
4

5 WHY EMBODIMENTS OF THE INVENTION FUNCTION:

6 In this embodiment the value counter 521 associated
7 with a route in the network route cache approximates the
8 relative value that route. The measured "value" is a
9 function of both the frequency and breadth of that route.
10 The subset H_i in which a particular route is placed by the
11 analyzer method 600, and the relative standing that route
12 will have within its subset H_i , will depend on both the
13 frequency and the breadth of that particular route. If,
14 for example, two (2) particular routes, R1 and R2, in a
15 particular subset H_i , have the same breadth value, but R1
16 is used more frequently (has a higher frequency value),
17 then R1 will have a greater value as measured by this
18 invention than route R2. Similarly, if two (2) particular
19 routes, R1 and R2 are used with equal frequency, but R1 is
20 used for a larger set of destinations (has greater
21 breadth), then R1 will have a greater value as measured by
22 this invention than R2.

23 These properties of the profiler can be better
24 understood with the following examples.

1 1.) Suppose the router is profiling two routes R1 and
2 R2 in a subset H_i . Further suppose that the
3 router is using route R1 to route 40,000 packets
4 per second evenly distributed between 5,000
5 unique destinations, additionally the router is
6 using route R2 to route 10,000 packets per second
7 evenly distributed between 5,000 unique
8 destinations. These two routes R1 and R2 then
9 have equal breadth but one of the two routes, R1,
10 is used more frequently than the other route, R2.
11 Given this traffic pattern then any time the
12 profiler is analyzed there will be more entries
13 in the profile table that are associated with
14 route R1 than with route R2. As a result after a
15 sampling period T_1 route R1 will have a greater
16 measured value than R2 as measured by evaluating
17 a sample from the profile table 220 in the
18 profiler 135. Given a random distribution of
19 packets and a suitably long sampling period T_1
20 the ratio of the number of entries associated
21 with R1 to the number of entries associated with
22 R2 found in the profiler table 220 will
23 approximately equal: $4/1 == 4$.

1 2.) Suppose the router is profiling two routes R1 and
2 R2 in a subset H_i . Further suppose that route R1
3 is used to route 10,000 packets per second evenly
4 distributed between 15,000 unique destinations,
5 and that route R2 is used to route 10,000 packets
6 per second evenly distributed between 5,000
7 unique destinations. These two routes then have
8 equal frequency of use but one route, route R1,
9 is used to route to a larger number of
10 destinations (has greater breadth) than the other
11 route R2. As a result route R1 will have a
12 greater measured value than R2 as measured by
13 evaluating a sample from the profile table 220 in
14 the profiler 135. Given a random distribution of
15 packets and a suitably long sampling period T_1
16 the ratio of the number of entries associated
17 with R1 to the number of entries associated with
18 R2 found in the profiler table 220 will approach
19 $3/1 == 3$ as the size of the profiler table 220
20 grows. (Typically the profiler table will be
21 fixed in size for a particular embodiment of this
22 invention, this example shows a relationship
23 between the selected size of the profiler table

1 and the sampling characteristics of an embodiment
2 of the invention.)
3

4 HOW EMBODIMENTS OF THE INVENTION CAN BE REVISED TO CHANGE
5 ITS PERFORMANCE CHARACTERISTICS:

6 Many of the characteristics the embodiments described
7 by figures 2, 5, and 6 can be varied in order to vary the
8 performance characteristics of these embodiments.

9 Given an embodiment of an analyzer 140 as described
10 above with Figures 4 and 5 the sampling period T1, the hash
11 function that is programmed into 230, and the subset count
12 N can all be varied to change the performance
13 characteristics of the embodiment.

14 The selected sampling period T1 has an effect on the
15 amount of influence breadth has on the measured value of
16 routes. As the sampling period T1 approaches 0, breadth of
17 use begins to influence the measured value of routes more
18 and more. When the sampling period T1 equals the minimum
19 amount of time, MTime, that will elapse between two packets
20 received by the routing path 115 and the number of subsets
21 N is 1 the destination for each and every packet routed
22 will appear in the profile table during analysis. When the
23 sampling period T1 equals MTime and the number of subsets N
24 is 1 the measured value becomes almost entirely a function

1 of breadth; with these configuration parameters frequency
2 has a much weaker influence on the measured value than
3 breadth. As the sampling time T_1 grows bigger than $MTime$
4 frequency of use begins to influence the measured value
5 more. Typically the sampling period T_1 will be much larger
6 than $MTime$, $MTime$ is used as an example to show how varying
7 T_1 varies the sampling characteristics of the profiler.

8 Additionally the subset count N may be varied to
9 increase the influence of breadth on the measured values.
10 When the profile table is saturated (nearly all of the
11 entries in the table have been written to in a sampling
12 period) each route in a subset H_i will have on average PTS
13 / HIS entries associated with them, where PTS is the number
14 of entries in the Profile Table 220 and HIS is the number
15 of routes in the subset H_i . This ratio PTS / HIS reflects
16 the minimum average number of destinations that are in use
17 within each route in a subset H_i , if that subset is
18 saturating the profile table during its sampling period.
19 Once the average number of hosts per route in a subset H_i
20 grows to exceed PTS/HIS the measured values for routes
21 within H_i becomes more a function of frequency than of
22 breadth. The subset count N may be increased to increase
23 the PTS/HIS value to increase the influence of breadth on
24 the measured route values.

1 The size of the profile table 220 in the profiler 200
2 also has an influence on the influence of breadth on the
3 measured values of routes in the network route cache 106.
4 The size of the table affects the PTS/HIS ratio that was
5 discussed above. Increasing the size of the profile table
6 220 increases the influence breadth has on the measured
7 value of routes.

8 And finally the hash function can be varied between
9 sampling periods in order to generate a more complete list
10 of all of the destinations that the router is sending
11 packets to.

12

13 ALTERNATE SOLUTIONS:

14 Three alternative solutions may be used but have
15 various problems. First, a large network route cache may
16 be used. The network route cache may have as many entries
17 as there are active routes in the FIB, in this case it
18 isn't necessary to profile the routes for inclusion in the
19 network route cache because in this case all of the active
20 entries are always present. The disadvantage of this
21 method is that it requires a very large and expensive high
22 performance network route table. Furthermore, since this
23 large network route table would most likely have to be off-
24 chip, the network route table would perform poorly compared

1 to a system that uses on-chip network route caches. A
2 route profiler is not necessary in this architecture
3 because every route in the host route cache has a
4 corresponding route in the high speed network route table.

5 Second, dirty bits may be used for each route in the
6 network route cache such that each time a network route
7 cache entry is used, the entry's dirty bit is cleared. The
8 route selection algorithm could then use these dirty bits
9 in an LRU replacement strategy. The disadvantage of this
10 method is that it only takes into account the frequency of
11 use for each route in the network route cache and does not
12 take into account the breadth of use for each network route
13 cache entry. As a result, the system in the router will
14 require a significantly larger host route cache because it
15 will use the network route cache entries inefficiently.

16 Third, the routes can be placed in the network route
17 cache whenever the system in the router receives a packet
18 that uses that route to get to its next hop destination.
19 When the network route cache is full, then randomly select
20 routes to be discarded from the network route cache. The
21 disadvantage of this method is that it does not consider
22 either the frequency of use or the breadth of use for any
23 of the routes in the network route cache. This will result
24 in an inefficient use of the network route cache and a

1 degradation of routing performance, and will require a
2 larger host route cache.

3

4 The various engines, tools, or modules discussed
5 herein may be, for example, software, firmware, commands,
6 data files, programs, code, instructions, or the like, and
7 may also include suitable mechanisms.

8

9 Reference throughout this specification to "one
10 embodiment", "an embodiment", or "a specific embodiment"
11 means that a particular feature, structure, or
12 characteristic described in connection with the embodiment
13 is included in at least one embodiment of the present
14 invention. Thus, the appearances of the phrases "in one
15 embodiment", "in an embodiment", or "in a specific
16 embodiment" in various places throughout this specification
17 are not necessarily all referring to the same embodiment.
18 Furthermore, the particular features, structures, or
19 characteristics may be combined in any suitable manner in
20 one or more embodiments.

21

22 Other variations and modifications of the above-
23 described embodiments and methods are possible in light of
24 the foregoing disclosure. Further, at least some of the

1 components of an embodiment of the invention may be
2 implemented by using a programmed general purpose digital
3 computer, by using application specific integrated
4 circuits, programmable logic devices, or field programmable
5 gate arrays, or by using a network of interconnected
6 components and circuits. Connections may be wired,
7 wireless, and the like.

8

9 It will also be appreciated that one or more of the
10 elements depicted in the drawings/figures can also be
11 implemented in a more separated or integrated manner, or
12 even removed or rendered as inoperable in certain cases, as
13 is useful in accordance with a particular application.

14

15 It is also within the scope of an embodiment of the
16 present invention to implement a program or code that can
17 be stored in a machine-readable medium to permit a computer
18 to perform any of the methods described above.

19

20 Additionally, the signal arrows in the
21 drawings/Figures are considered as exemplary and are not
22 limiting, unless otherwise specifically noted.
23 Furthermore, the term "or" as used in this disclosure is
24 generally intended to mean "and/or" unless otherwise

1 indicated. Combinations of components or steps will also
2 be considered as being noted, where terminology is foreseen
3 as rendering the ability to separate or combine is unclear.
4

5 As used in the description herein and throughout the
6 claims that follow, "a", "an", and "the" includes plural
7 references unless the context clearly dictates otherwise.
8 Also, as used in the description herein and throughout the
9 claims that follow, the meaning of "in" includes "in" and
10 "on" unless the context clearly dictates otherwise.
11

12 It is also noted that the various functions,
13 variables, or other parameters shown in the drawings and
14 discussed in the text have been given particular names for
15 purposes of identification. However, the function names,
16 variable names, or other parameter names are only provided
17 as some possible examples to identify the functions,
18 variables, or other parameters. Other function names,
19 variable names, or parameter names may be used to identify
20 the functions, variables, or parameters shown in the
21 drawings and discussed in the text.
22

23 While the present invention has been described herein
24 with reference to particular embodiments thereof, a

1 latitude of modification, various changes and substitutions
2 are intended in the foregoing disclosures, and it will be
3 appreciated that in some instances some features of the
4 invention will be employed without a corresponding use of
5 other features without departing from the scope and spirit
6 of the invention as set forth. Therefore, many
7 modifications may be made to adapt a particular situation
8 or material to the essential scope and spirit of the
9 present invention. It is intended that the invention not
10 be limited to the particular embodiment disclosed as the
11 best mode contemplated for carrying out this invention, but
12 that the invention will include all embodiments and
13 equivalents falling within the scope of the appended
14 claims.